

Central auditory system plasticity: Generalization to novel stimuli following listening training

Kelly Tremblay^{a)}

Department of Communication Sciences and Disorders, Auditory Neuroscience Laboratory,
Northwestern University, 2299 N. Campus Drive, Evanston, Illinois 60208-3540

Nina Kraus

Department of Communication Sciences and Disorders, Auditory Neuroscience Laboratory, Northwestern
University, 2299 N. Campus Drive, Evanston, Illinois 60208-3540 and Departments of Neurobiology
and Physiology, and Otolaryngology, Northwestern University, Chicago, Illinois 60208-3540

Thomas D. Carrell

Department of Communication Disorders, University of Nebraska, Lincoln, Nebraska 68588

Therese McGee

Department of Communication Sciences and Disorders, Auditory Neuroscience Laboratory,
Northwestern University, 2299 N. Campus Drive, Evanston, Illinois 60208-3540

(Received 3 June 1996; revised 17 June 1997; accepted 5 August 1997)

Behavioral perceptual abilities and neurophysiologic changes observed after listening training can generalize to other stimuli not used in the training paradigm, thereby demonstrating behavioral “transfer of learning” and plasticity in underlying physiologic processes. Nine normal-hearing monolingual English-speaking adults were trained to identify a prevoiced labial stop sound (one that is not used phonemically in the English language). After training, the subjects were asked to discriminate and identify a prevoiced alveolar stop. Mismatch negativity cortical evoked responses (MMN) were recorded to both labial and alveolar stimuli before and after training. Behavioral performance and MMNs also were evaluated in an age-matched control group that did not receive training. Listening training improved the experimental group’s ability to discriminate and identify an unfamiliar VOT contrast. That enhanced ability transferred from one place of articulation (labial) to another (alveolar). The behavioral training effects were reflected in the MMN, which showed an increase in duration and area when elicited by the training stimuli as well as a decrease in onset latency when elicited by the transfer stimuli. Interestingly, changes in the MMN were largest over the left hemisphere. The results demonstrate that training can generalize to listening situations beyond those used in training sessions, and that the preattentive central neurophysiology underlying perceptual learning are altered through auditory training. © 1997 Acoustical Society of America. [S0001-4966(97)00912-0]

PACS numbers: 43.71.Pc [WS]

INTRODUCTION

The auditory cortex is plastic, that is, it is capable of reorganization as a function of experience. This fact has been well documented by animal experiments, and by human behavioral and electrophysiologic studies. For example, single- and multiple-unit recordings from mammalian cortical neurons show learning-induced changes in the auditory cortex (Buchwald *et al.*, 1966; Olds *et al.*, 1972; Kraus and Disterhoft, 1982; Diamond and Weinberger, 1984, 1986, 1989; Bakin and Weinberger, 1990; Recanzone *et al.*, 1993; Edeleine *et al.*, 1993; Weinberger, 1993). In humans, Näätänen *et al.* (1993) have reported learning-related neurophysiologic changes reflected in a far-field electrophysiologic response called the mismatch negativity (MMN). They showed that the MMN changed following discrimination training using

tonal stimuli. Similarly, Kraus *et al.* (1995a) demonstrated learning-related changes in the MMN when elicited by synthetic speech stimuli.

The MMN is an auditory-evoked response that reflects preattentive discrimination of acoustic signals (Näätänen *et al.*, 1978; Sams *et al.*, 1985; Näätänen and Picton, 1987; Scherg and Picton, 1990; Giard *et al.*, 1990; Csépe *et al.*, 1987). The MMN is elicited by a rarely occurring stimulus when it is presented within a series of homogeneous stimuli (Näätänen *et al.*, 1978). Animal and human studies indicate that the encoding of different acoustic cues important for speech perception involve distinct neuroanatomic and neurophysiologic processes (Blackburn and Sachs, 1990; Delgutte and Kiang, 1984; Steinschneider *et al.*, 1982, 1990) and that the MMN can be used to probe the neurophysiologic encoding of essential speech features (McGee *et al.*, 1996; Micco *et al.*, 1995; Csépe, 1995; Sharma *et al.*, 1995; Kraus *et al.*, 1993, 1994a, 1994b, 1995b, 1996).

The Näätänen *et al.* (1993) and Kraus *et al.* (1995a) findings suggest that changes in the MMN reflect experience-related changes in central auditory processing of

^{a)}Present address: House Ear Inst., Electrophysiology Lab., 2100 West Third St., Los Angeles, CA 90069.

complex signals such as speech. However, in those studies, subjects were trained and tested using the same stimuli. What is of great interest is if changes in stimulus representation are limited to the stimuli used in training, or whether the changes in acoustic representation at a preattentive level will generalize to novel stimuli with acoustic properties similar to the trained stimuli. This “transfer of learning” is important from a clinical perspective because it is important to know if efforts at (re)habilitation using auditory training will generalize beyond the training sessions to other listening events.

Behavioral evidence of “transfer of learning” already has been established in the speech-learning literature. For example, McClaskey *et al.* (1983) found that adult listeners generalized their newly acquired ability to perceive prevoiced labial syllables to prevoiced alveolar phonemes. However, little is known about the neurophysiologic changes associated with such learning and transfer of learning.

Consequently, this study combined the behavioral approaches of Pisoni *et al.* (1982) and McClaskey *et al.* (1983) with the neurophysiological approach used by Nääätänen *et al.* (1993) and Kraus *et al.* (1995a) to explore whether behavioral training modified the neurophysiologic representation of speech, and to test the hypothesis that neurophysiologic modification is not limited to the stimuli used to establish the change. Specifically, this study was designed to determine if surface-recorded mismatch negativity responses in humans reflect behavioral changes in perception following training, and whether “transfer of learning” is evident behaviorally and neurophysiologically in response to novel speech contrasts having an equivalent voice-onset-time distinction but a different place of articulation.

I. METHODS

A. Subjects

The subjects were 18 normal hearing, right-handed monolingual speakers of English. Normal hearing was defined as pure tone sensitivity better than or equal to 20 dB HL bilaterally for octave frequencies between 250 Hz and 8 kHz. All subjects were born into monolingual English-speaking households and had less than five years of exposure to a foreign language(s) in school. Each subject was randomly assigned to either an experimental or control group. The mean age for the control group was 21 years with an age range of 18–28 years. The mean age for the experimental group was 22 years with an age range of 18–26 years.

B. Stimuli

The stimuli were synthesized tokens modeled after those used by McClaskey *et al.* (1983). Labial (/ba/–/pa/) and alveolar (/da/–/ta/) continua were created in which voice-onset-time (VOT) varied from –50 ms to +50 ms in 10-ms steps. Therefore, each continuum consisted of eleven items. The VOT is defined as the interval between the release from stop closure and the onset of laryngeal pulsing. A negative VOT indicates laryngeal pulsing during the stop closure period before the release. The stimuli were generated using a Klatt digital speech synthesizer (Klatt, 1980).

The steady-state portion of the stimuli consisted of the vowel /a/, which varied in duration relative to the VOT so that the overall duration for each stimulus remained constant at 180 ms. The formant values for this vowel were: $F1 = 700$ Hz, $BW1 = 90$ Hz; $F2 = 1200$ Hz, $BW2 = 90$ Hz; $F3 = 2600$ Hz, $BW3 = 130$ Hz, $F4 = 3300$ Hz, $BW4 = 400$ Hz; $F5 = 3700$ Hz, $BW5 = 500$ Hz. The fundamental frequency of the stimuli began at 120 Hz and then fell to 100 Hz during the steady-state portion of the vowel. The labial formant transitions were 40 ms in duration. The alveolar formant transitions were 50 ms in duration. To simulate a burst, a turbulent noise source (AF) 10 ms in duration and 60 dB in amplitude was added to both bilabial and alveolar stimuli. The spectrum of the burst was centered around 2500–4000 Hz.

Because the MMN is elicited by any acoustic change, we chose to use synthetic rather than natural speech in order to precisely control the acoustic parameters and to determine systematically which acoustic dimension was reflected in the MMN. The use of natural speech would have complicated the interpretation because natural speech contains many simultaneously changing parameters that might be reflected in the MMN.

C. Procedure

The entire testing procedure took nine days for each individual. The experimental and control group participated in exactly the same testing on days 1, 2, 8, and 9. The only difference was that the experimental group participated in training sessions on days 3–7 whereas the control group did not receive any training.

1. Day 1

a. Electrophysiology baseline. Baseline MMN responses were obtained on the first day for both the experimental and control groups in response to the two stimulus pairs; the training stimuli (labial) and the transfer stimuli (alveolar). The –10-ms VOT stimulus (either labial or alveolar) was the standard stimulus (frequency of occurrence = 85%) and the –20-ms VOT stimulus (either labial or alveolar) was the deviant stimulus (frequency of occurrence = 15%). The stimuli were presented in a pseudorandom sequence with at least three standard stimuli separating presentations of deviant stimuli (oddball paradigm). Although 3500 standard stimuli were presented, the responses to the standard stimuli following each deviant were not included in the averaged response. In an “alone” condition, 1000 presentations of the deviant stimulus were presented alone for comparison with the response to that same stimulus when it was the deviant stimulus in the oddball paradigm. The rationale for using a deviant-stimulus-alone control condition is reviewed by Kraus *et al.* (1995b). The order of labial and alveolar presentation was counterbalanced across subjects to prevent any potential order effects. For each individual, the order of stimulus presentation was the same before and after training.

All subjects were tested in a sound-treated booth. The stimuli were presented monaurally to the right ear using Etymotic Research (ER3) insert earphones. Subjects were instructed to ignore the stimuli. To ensure that the subject was

not attending to the stimuli, each subject watched a videotape of his or her choice. The test stimuli were presented at 74 dB SPL while the background videotape sound, presented in free field, was 40 dB SPL.

Previous studies by Näätänen *et al.* (1993) and Kraus *et al.* (1995a) recorded over the midline of the scalp only and therefore did not report topographical distributions of MMN change. This study included nine active electrode locations. Active electrodes were placed at Fz, Cz, Pz, over the frontal lobes (*F3* and *F4*), temporal lobes (TL and TR), and mastoids (A1 and A2). Electrode TR was situated halfway between T4 and T6. Electrode TL was situated half way between T3 and T5 according to the International 10/20 recording system (Jasper, 1958). A nose electrode served as the reference and a forehead electrode served as ground. Eyeblink activity was monitored using electrodes located on the superior and outer canthus of one eye. Artifacts measuring in excess of 100 μV on the EOG channel were rejected off-line during data processing.

Each individual's grand-averaged "difference" wave was calculated by subtracting the average response to the deviant stimulus when presented alone from the response to the same stimulus when it was presented in the oddball paradigm. The MMN onset latency, duration, and area were measured from the individual subject grand-averaged difference wave. Latency was measured between 120 and 500 ms. Duration was defined as the offset minus the onset latency. The MMN area was calculated by drawing a line between the onset and offset of the MMN in the difference wave, and determining the number of $\text{ms} \times \mu\text{V}$.

Individual grand averages were averaged together for each group. Group MMN onset latency, duration, and area measurements were calculated using the same methods. The present study intentionally focused on group effects. To date, methods for statistically defining the MMN in individuals are evolving (McGee, 1997a, 1997b). Extracting the MMN signal from noise in individual responses has been a longstanding methodological issue.

2. Day 2

a. Behavioral baseline: Discrimination testing. During behavioral tests, subjects were seated in a sound-treated booth in front of a computer monitor which provided visual instructions to the subjects. Subjects in both the experimental and control groups were instructed to listen to two stimuli and to indicate whether the stimuli were the same or different. This AX discrimination task served as a practice task and included ten trials without feedback using two-step stimulus pairs from the labial series. Once the subjects demonstrated proficiency with this task, they performed the same discrimination test using all of the labial stimuli in the VOT continuum. That is, each subject was presented with 264 trials consisting of both one- and two-step pairs of the labial stimuli with no feedback. The two-step pairs were included to provide reinforcement since these were "easier" contrasts.

The discrimination task was an AX paradigm with an interstimulus interval of 750 ms. The interpair interval varied since the test was self-paced. Once the subject provided a

response, the next stimulus pair was delivered 1 s later. The discrimination test was repeated using stimuli from the alveolar continuum. The number of correct and incorrect responses for the $-20/-10$ ms VOT contrast were calculated for each individual. These results served as a baseline discrimination measure. The order of labial and alveolar presentation was counterbalanced across subjects to prevent influences of order effects. For each individual, the order of stimulus presentation was the same before and after training.

b. Identification testing. Following discrimination testing, each subject was asked to identify, from three choices displayed as text on the computer screen, what sound they heard. This paradigm was a three-alternative forced-choice identification task. When a single labial stimulus from the labial continuum was presented, the choices were /mba/, /ba/, and /pa/. When the stimulus was from the alveolar continuum, the choices were /nda/, /da/, and /ta/. Each test consisted of 210 trials and included all 11 items from the continuum. No feedback was provided to the subject. However, in order to score identification data, labels were arbitrarily assigned such that the $-50-$, $-40-$, $-30-$, and $-20-$ ms VOT stimuli were assigned the label /mba/, the $-10-$, $0-$, $+10-$, and $+20-$ ms VOT were assigned the label /ba/, and the $+30-$, $+40-$, and $+50-$ ms VOT stimuli were assigned the label /pa/. Percent correct was calculated for the number of times the $-20-$ ms VOT stimulus was identified as /mba/. These results served as a baseline identification measure.

3. Day 3

a. Behavioral training (experimental group only). The experimental subjects were asked to listen to single presentations of the labial stimuli in order from $-50-$ ms through $+50-$ ms VOT. This served as a familiarization session. Using a fading technique described by Jamieson and Morosan (1986, 1989), subjects were presented with either a $-10-$ ms or $-30-$ ms VOT stimulus from the labial continuum. Subjects were asked to identify the sound as either /mba/ or /ba/. Both choices were presented as text on the computer monitor. Feedback in the form of a green reinforcement light appeared when the subject correctly identified the $-30-$ ms VOT stimuli as /mba/ and the $-10-$ ms VOT stimuli as /ba/. This task was repeated 50 times to train the listener to label the acoustic prevoiced component as /mba/.

The $-10-$ ms and $-30-$ ms stimuli were chosen based on a pilot study in which listeners were able to identify these stimuli well above chance without training. Therefore, this initial session allowed the subjects to listen to the prevoiced stimuli and orient themselves to the prevoiced cue using an easy stimulus pair. Following this fading session, each subject began their first training session using a more difficult stimulus ($-20-$ ms VOT). This stimulus was now assigned the label /mba/. The training session consisted of four blocks of 50 trials where either a $-10-$ ms or $-20-$ ms VOT labial stimulus was presented. Feedback in the form of a green reinforcement light appeared when the $-20-$ ms VOT stimuli was labeled as /mba/ and the $-10-$ ms VOT was labeled as /ba/. Each stimulus was presented randomly with equal probability of occurrence. Percent correct was calculated based

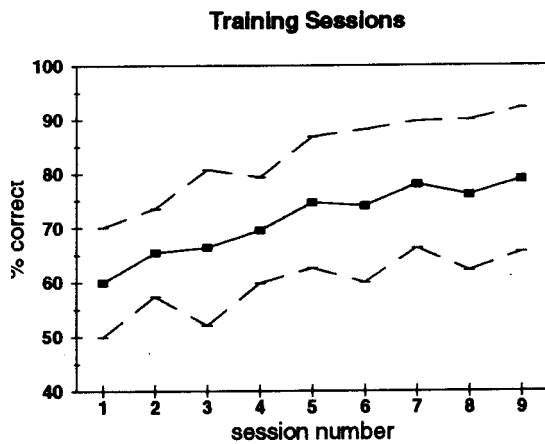


FIG. 1. Average correct identification scores (■---■) and one standard deviation (---) are shown as a function of training session for the experimental group. Identification scores improved significantly from the first to final training session ($n=9$).

on the number of correct responses for each block of 50 trials. Performance on all four blocks of 50 trials was averaged to obtain a final training session score.

4. Days 4–7

a. Behavioral training continued (experimental group only). Each subject in the experimental group continued the training sessions, listening, and identifying the -20 - and -10 -ms VOT labial stimuli. Nine sessions were conducted over 5 days. Each training session (200 trials) lasted approximately 20 min.

5. Day 8

a. Electrophysiology posttraining test. Electrophysiologic responses were measured for both groups in response to the -20 - and -10 -ms VOT contrasts (labial and alveolar) using the same procedure described for day 1.

6. Day 9

a. Behavioral posttraining test. Discrimination and identification measures were obtained for both the labial and alveolar stimuli for both the experimental and control groups as described under day 2.

II. RESULTS

The results are presented in two sections: (A) Behavioral data and (B) Electrophysiological data.

A. Behavioral results

1. Identification training

Through training, all subjects in the experimental group learned to identify the -20 -ms VOT stimulus as /mba/ and the -10 -ms VOT stimulus as /ba/. Grouped data show a significant improvement from the first to the last training session ($t=3.43$, $df=8$, $p<0.01$). Group averaged data across sessions are shown in Fig. 1.

2. Pre- and postdiscrimination and identification testing

a. Trained stimuli /mba/. Training resulted in improved posttraining discrimination and identification scores for the experimental group but not the control group (Fig. 2). Planned comparisons revealed a significant improvement in the experimental group's ability to identify the -20 -ms VOT stimuli as /mba/ when comparing pre- versus posttraining identification scores ($t=2.11$, $df=8$, $p<0.05$). The control group did not show any significant change in performance from pre- to posttesting ($t=0.85$, $df=8$, $p>0.05$). Similarly, the experimental group demonstrated a significant improvement in their ability to discriminate the -20 -ms VOT /mba/ contrast ($t=2.39$, $df=8$, $p<0.05$) whereas the control group did not ($t=1.51$, $df=8$, $p>0.05$).

b. Transfer stimuli. The effects of training transferred to the /nda/ stimuli with posttraining discrimination scores improving significantly for the experimental group ($t=2.60$, $df=8$, $p>0.05$) but not the control group ($t=1.71$, $df=8$, $p>0.05$). Improvement in the identification of the -20 -ms VOT /nda/ stimulus narrowly missed reaching significance for the experimental group ($t=1.71$, $df=8$, $p=0.06$). The control group did not show any significant changes in the ability to correctly identify the -20 -ms VOT /nda/ stimulus ($t=0.09$, $df=8$, $p>0.05$).

3. Summary

In summary, listening training improved the experimental group's ability to discriminate and identify the unfamiliar prevoiced labial stimulus /mba/. That enhanced ability transferred from one place of articulation (labial) to another (alveolar). No significant changes in performance were seen for the control group since they did not receive the training.

B. Electrophysiology results

1. MMN onset latency, duration, and area

Planned comparisons revealed a significant decrease in MMN onset latency for the experimental group but not the control group for the training stimuli /mba/ ($t=2.31$, $df=8$, $p>0.05$; $t=0.99$, $df=8$, $p>0.05$). No significant changes were seen in onset latency for either the experimental or control groups when tested using the transfer stimuli ($t=1.48$, $df=8$, $p>0.05$; $t=0.22$, $df=8$, $p>0.05$).

Planned comparisons comparing pre- versus posttraining scores revealed significant increases in MMN duration for the experimental group for both the trained and transfer stimuli ($t=4.97$, $df=8$, $p>0.01$; $t=2.91$, $df=8$, $p<0.05$). The control group did not show significant changes for either stimulus condition ($t=0.90$, $df=8$, $p>0.05$; $t=0.25$, $df=8$, $p>0.05$).

Planned comparisons comparing pre- versus posttraining revealed significant increases in MMN area for the experimental group for both the trained and transfer stimuli ($t=2.77$, $df=8$, $p>0.05$; $t=2.94$, $df=8$, $p>0.05$). The control group did not show significant changes for either stimulus condition ($t=1.13$, $df=8$, $p<0.05$; $t=0.91$, $df=8$, $p>0.05$).

Identification of -20 msec VOT

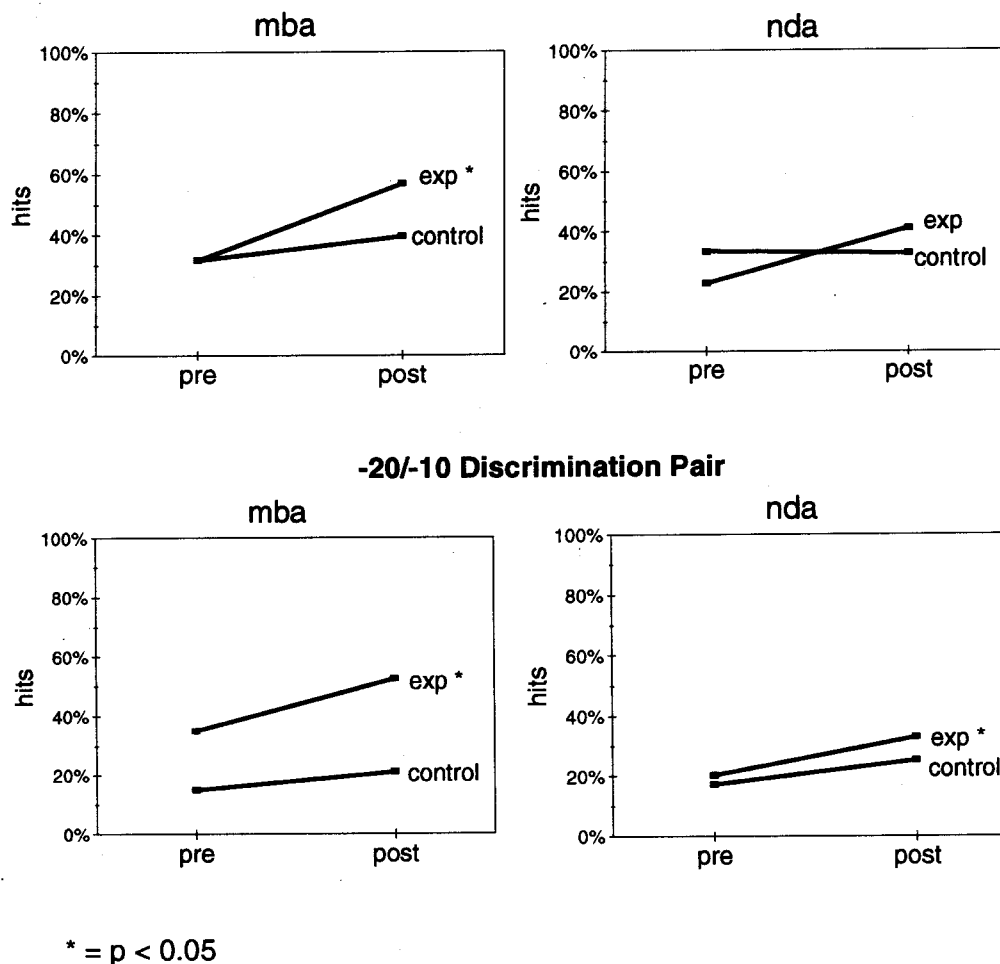


FIG. 2. Pre- versus posttraining identification and discrimination performance for the training and transfer stimuli for the experimental and control groups. Asterixes denote a significant change in performance with p values < 0.05 .

In summary, an increase in MMN duration and area were seen for the experimental group when elicited by either the training or transfer stimuli following training (Fig. 3). No significant pre- versus posttraining changes were seen for the control group.

2. Laterality differences

Changes in MMN were largest over the left hemisphere. Figure 4 provides right versus left hemisphere MMN data pre- and posttraining for the experimental group in response to the training stimuli. Enhancement of the MMN was evident visually at all electrode sites in response to the training stimuli. For the transfer stimuli, changes were evident at Fz and FL only. Because there appeared to be a greater amount of change over the left hemisphere in the experimental group, MMN onset latency, duration, and area measures were analyzed using a $2 \times (2 \times 2)$ repeated measures ANOVA (experimental versus control group, left versus right hemisphere, and frontal versus temporal lobe), using post- minus pretraining values as the dependent variable.

For the training stimuli (mba), there were significant group effects for onset latency ($F=4.135$, $p=0.05$). A

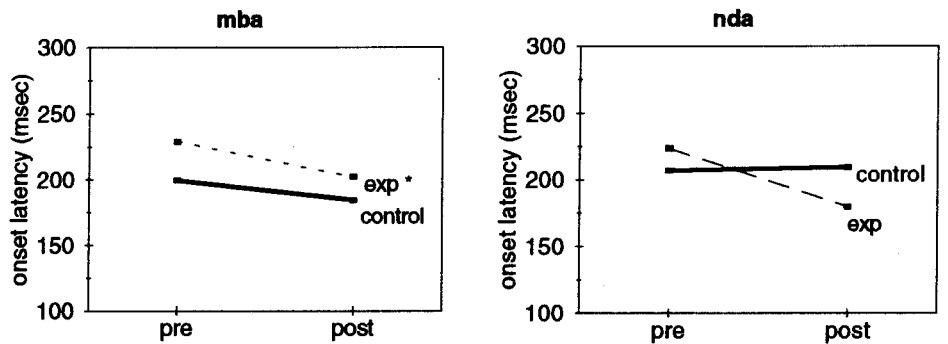
group \times hemisphere interaction was seen for MMN onset latency ($F=5.62$, $p<0.05$). *Post hoc t* tests revealed that the significant difference between right and left onset latency was seen over the frontal lobes. The decrease in onset latency following training was greater at electrode site FL when compared to FR ($t=3.10$, $df=8$, $p<0.01$).

For the training stimuli (mba), there were significant group effects for MMN duration and area ($F=4.41$, $p<0.05$; $F=5.63$, $p<0.05$). A group \times hemisphere \times lobe interaction was seen for MMN duration ($F=4.6$, $p<0.05$). *Post hoc t* tests revealed that the degree of duration change was greater at electrode site FL compared to FR in the experimental group ($t=3.74$, $df=8$, $p<0.01$). No significant onset latency, duration, or area differences were seen across hemispheres for the control group. For the transfer stimuli (nda), there were no significant hemispheric differences.

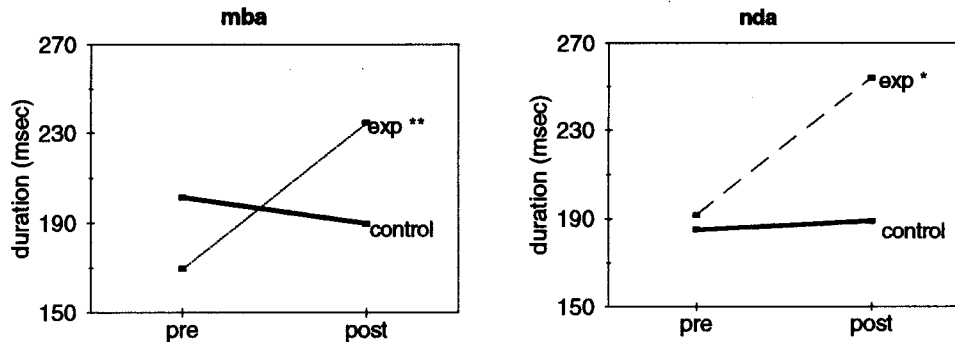
C. Summary of results

Listening training improved the ability to discriminate and identify an unfamiliar VOT contrast. That enhanced ability transferred from one place of articulation (labial) to another (alveolar). The behavioral training effects were re-

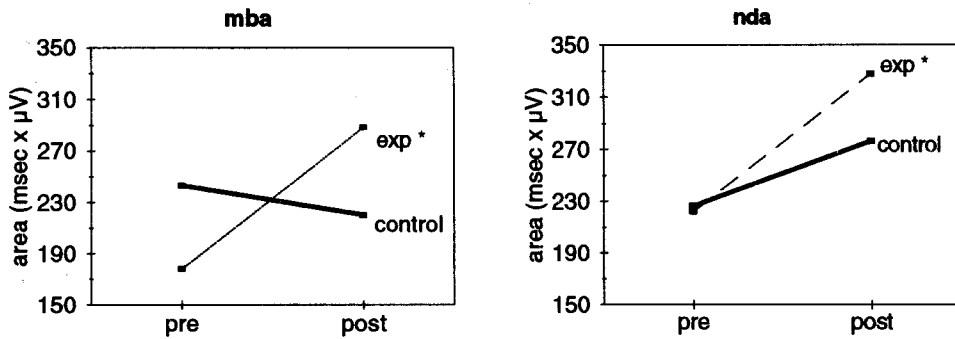
ONSET LATENCY



DURATION



AREA



* = $p < 0.05$
 ** = $p < 0.01$

FIG. 3. Increases in MMN duration and area and a decrease in onset latency (recorded from Fz) were seen in the experimental group and not in the control group for both training and transfer stimuli. Single asterixes denote a significant change with p values < 0.05 . Double asterixes denote a significant change with p values < 0.01 .

flected in the MMN, which showed an increase in duration and area when elicited by the training or transfer stimuli for the experimental group but not the control group. This finding is illustrated in Figs. 5 and 6, which show grand-averaged MMN waveforms for each group. Figure 7 illustrates that the grand-averaged MMN measures reflect the individual trends in the data and are not simply a result of one or two outliers. Figure 4 shows that changes in the MMN were largest over the left hemisphere.

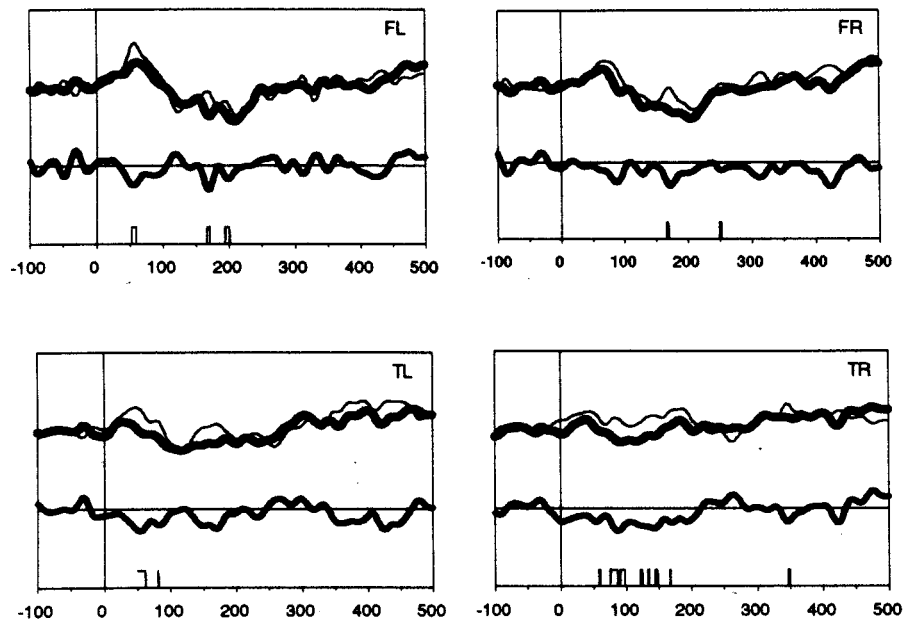
III. DISCUSSION

A. Main findings

Auditory training alters the perception of novel stimuli. Behavioral learning was shown by the improved identifica-

tion scores during training sessions as well as the improved pre- versus posttraining discrimination and identification scores. Concurrently, neurophysiological plasticity was shown by the changes in the mismatch response. Specifically, this study showed that (1) through training, the experimental group learned to identify speech contrasts (differing in VOT) that are not phonemically different in the English language, (2) improvement of voice-onset-time discrimination generalized to spectrally different stimuli that also varied in VOT, (3) electrophysiological responses to the training and transfer stimuli reflected this improved perception as an increase in MMN duration and area, as well as a decrease in the onset of the mismatch response for the experimental group but not the control group, and (4) electrophysiological

Pre-Training Condition



Post-Training Condition

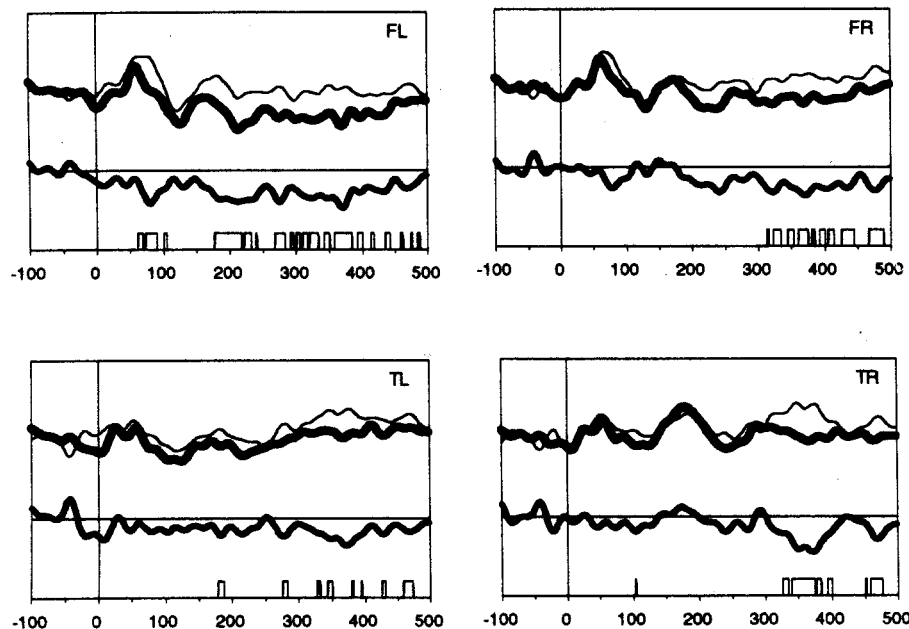


FIG. 4. Pre- and posttraining group-averaged electrophysiologic responses at four electrode sites (FL, TL, FR, and TR) are shown for the experimental group to the training stimuli ($-20/-10$ -ms VOT) when the deviant was presented alone (thin line) and when it was presented in the mismatch condition (thick line). The MMN is seen in the subtraction wave below. The area of significance ($p < 0.05$) is depicted by the blocked area on the x axis ($n = 9$).

duration changes were greater over the left hemisphere than the right hemisphere.

The behavioral data are consistent with those of McClaskey *et al.* (1983) who reported that the perception of a new linguistic contrast can be acquired by adult listeners and that this trained ability generalizes to novel stimuli not used during training. The electrophysiologic data reinforce the

findings of Näätänen *et al.* (1993) and Kraus *et al.* (1995a) that neurophysiologic changes can be demonstrated in humans following listening training, thereby reflecting auditory system plasticity. Most importantly, this study showed that both behavioral and electrophysiologic changes are generalizable to other auditory stimuli that were not used during the training sessions.

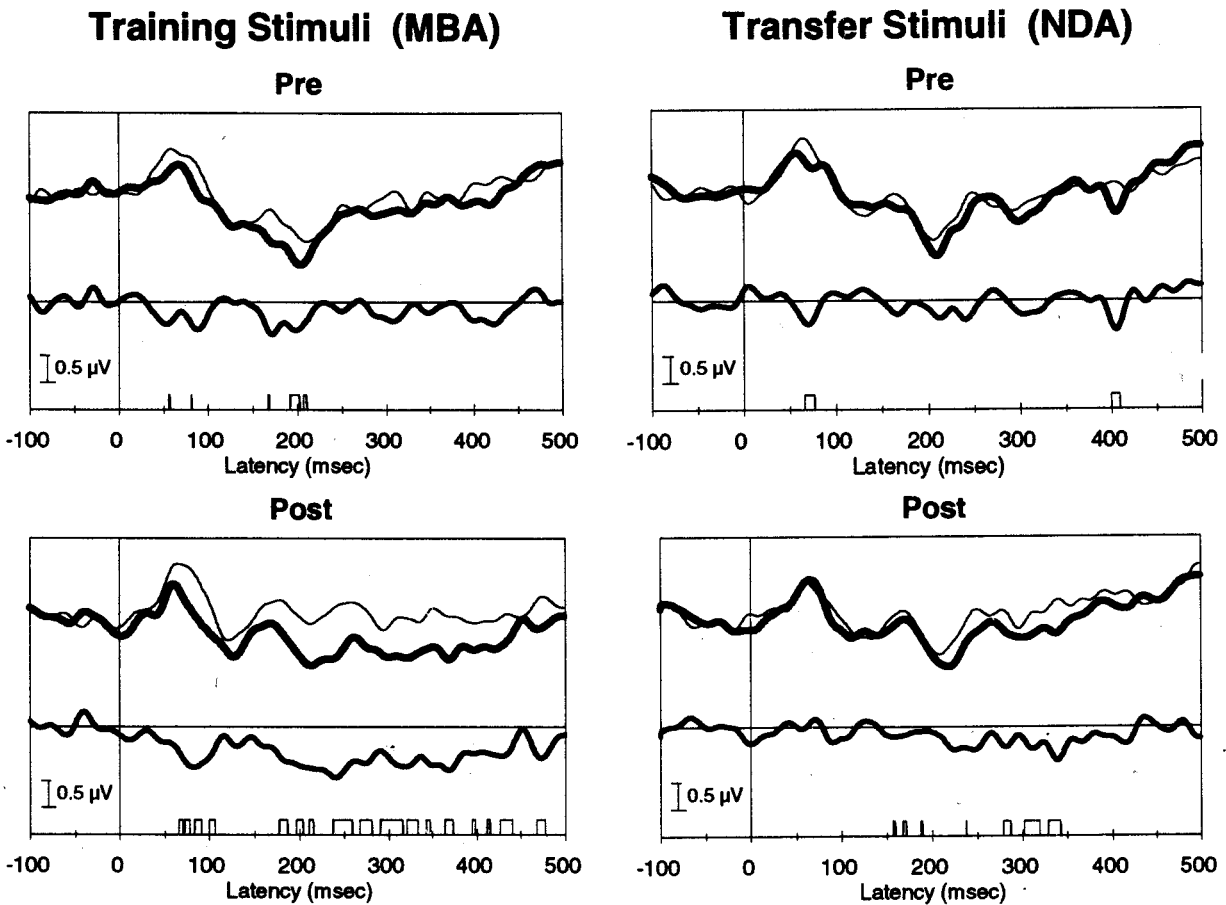


FIG. 5. A significant increase in duration and area of the mismatch response was seen following training for the experimental group. Group-averaged electrophysiological responses (electrode Fz) are shown for the experimental group to the training and transfer stimuli ($-20/-10$ -ms VOT) when the deviant stimulus was presented alone (thin line) and when it was in the mismatch paradigm (thick line). The MMN is seen in the subtraction wave below. The area of significance ($p < 0.05$) is depicted by the blocked area on the x axis ($n = 9$).

B. Implications of experimental results to other areas of research

1. Plasticity of the adult system

In the past, there was debate concerning the flexibility of the mature perceptual system to accommodate new phonetic categories. Both successful as well as unsuccessful training and generalization attempts subsequently have led to the conclusion that the adult system does not lose the ability to acquire new linguistic contrasts, provided specific experimental procedures are used to train the listeners (Carney *et al.*, 1977; Aslin and Pisoni, 1980; Strange and Dittmann, 1984; Jamieson and Morosan, 1989; Pisoni *et al.*, 1982; Werker and Tees, 1984a; Lively *et al.*, 1993, 1994; Logan *et al.*, 1991). Until now, this line of research has been limited to behavioral studies. The present study establishes that the adult perceptual system is indeed plastic and that transfer of training can be shown through both behavioral and neurophysiologic measures.

2. Locus of neurophysiologic change

Neurophysiological evidence of learning-induced plasticity as well as representation of VOT have both cortical and subcortical origins (Disterhoft and Stuart, 1976, 1977; Ryugo

and Weinberger, 1978; Kraus and Disterhoft, 1982; Weinberger *et al.*, 1984; Edeline *et al.*, 1990; Sinex and McDonald, 1988; Steinschneider *et al.*, 1982, 1990, 1995; Eggemont, 1995; McGee *et al.*, 1996). Furthermore, there is evidence for cortical and subcortical mismatch generators (Näätänen *et al.*, 1978, 1980; Hari *et al.*, 1984; Csépe *et al.*, 1987; Näätänen and Picton, 1987; Giard *et al.*, 1990; Scherg and Picton, 1990; Sams and Näätänen, 1991; Alho *et al.*, 1992; Javitt *et al.*, 1992; Kraus *et al.*, 1994a, 1994b). Training-induced neurophysiologic changes as well as their generalizability may be the result of (1) a greater number of neurons responding in the sensory field, (2) increased synchrony of neuronal responses, or (3) an expansion of receptive fields (Berger *et al.*, 1976; Merzinech *et al.*, 1983, 1984; Kaas *et al.*, 1983; Merzinech and Jenkins, 1983).

Although this study does not specifically address whether changes occurred cortically or subcortically, it does provide insight into where neural changes might have occurred and how neuroanatomy may relate to perceptual learning. Specifically, changes in MMN were greater over the left compared to the right hemisphere following training, thereby suggesting that preattentive aspects of speech processing may also be lateralized. This laterality finding is new and may have implications regarding the role of each hemi-

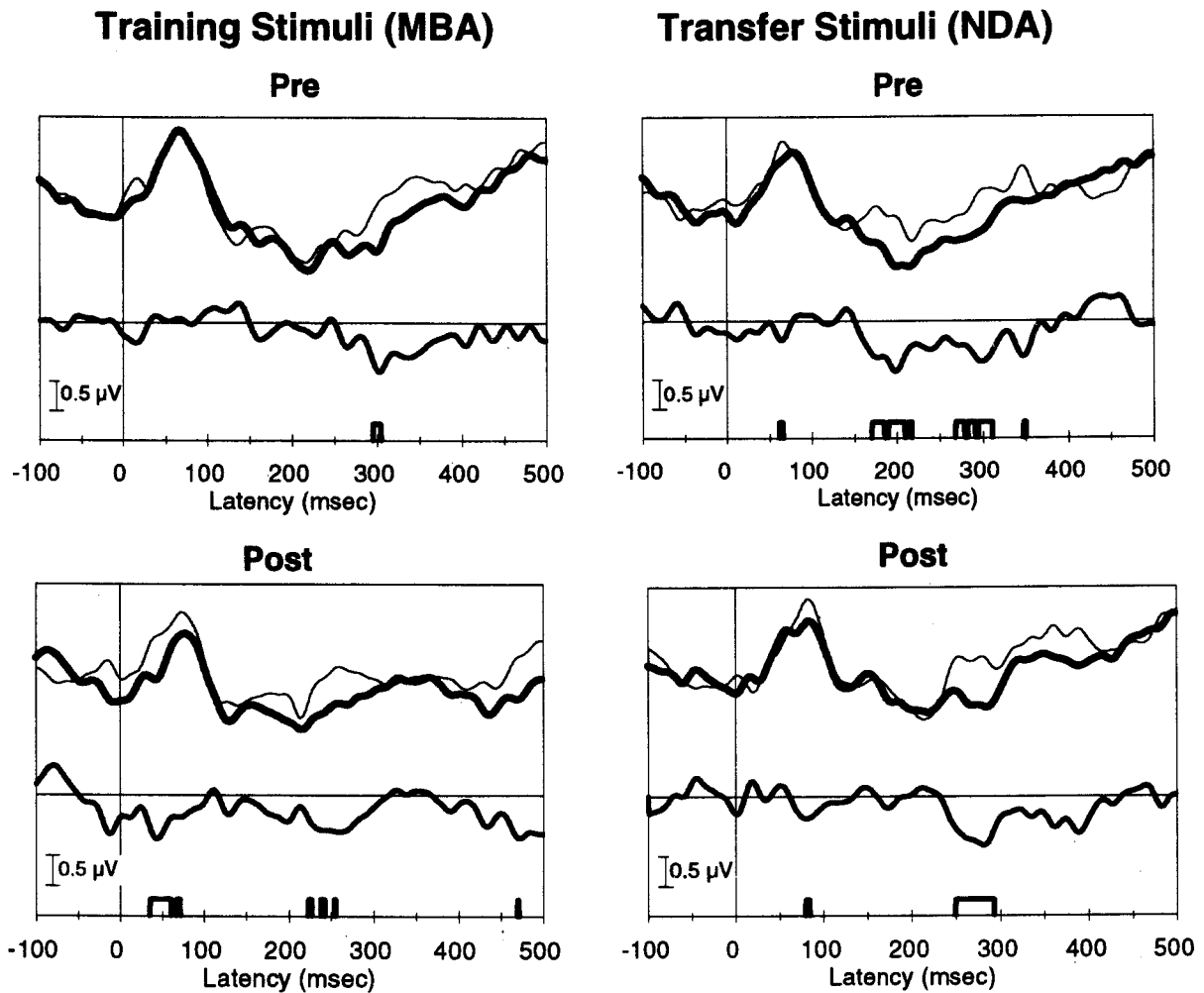


FIG. 6. No significant changes in onset latency, duration, and area of the mismatch response was seen for the control group. Group-averaged electrophysiologic responses (electrode Fz) are shown for the control group to the training and transfer stimuli ($-20/-10$ -ms VOT) when the deviant stimulus was presented alone (thin line) and when it was in the mismatch paradigm (thick line). The MMN is seen in the subtraction wave below. The area of significance ($p < 0.05$) is depicted by the blocked area on the x axis ($n = 9$).

sphere in processing learning-associated stimulus change.

Mapping studies using current source density analysis have reported both bilateral cortex involvement and predominantly right-hemispheric frontal involvement in MMN generation (Giard *et al.*, 1990; Paavilainen *et al.*, 1991). In the present study, the pretraining MMN was seen bilaterally and was not significantly larger over one hemisphere. The absence of laterality may result in part from the nature of the stimuli. Prior to training, subjects did not perceive the /mba/ and /nda/ stimuli to be phonetically distinct from /ba/ and /da/. Identification training may have resulted in these stimuli being recognized linguistically rather than as simple acoustic variants of /ba/ and /da/, thereby leading to greater activity in the left hemisphere. For example, when Giard *et al.* (1990) and Paavilainen *et al.* (1991) observed right-hemispheric dominance when recording the MMN, they used tonal stimuli. Because tones do not convey linguistic or phonetic information, the left hemisphere may not have played a dominant role in encoding those stimuli. Sharma and Kraus (1995) showed that the MMN lateralizes to the left if the change in a stimulus sequence signals a phonetic rather than a tonal pitch contrast. Thus, these studies in combination

with the findings from the present study suggest that the linguistic/phonetic nature of the stimuli may relate to which hemisphere is most active in detecting the stimulus change and that preattentive aspects of speech processing are modifiable and may also be lateralized.

3. The role of attention and perceptual learning

Regardless of where and how these neurophysiological changes take place, the data suggest that some degree of speech learning occurs at a preattentive level that can be measured in the absence of an overt response. Until now, the role of attention in comparable training and generalization experiments has been explored primarily in the visual modality. Visual experiments suggest that high-level attentional mechanisms are necessary for demonstrating "transfer" of visual learning (Ahissar and Hochstein, 1993). Even though the behavioral tasks in this study required attention and cognitive processing of the stimuli during training, changes in the MMN elicited by the training as well as the transfer stimuli indicate that speech-sound learning is evident at an automatic, preattentive level. Future studies examining the role of attention and other high-level cognitive processes

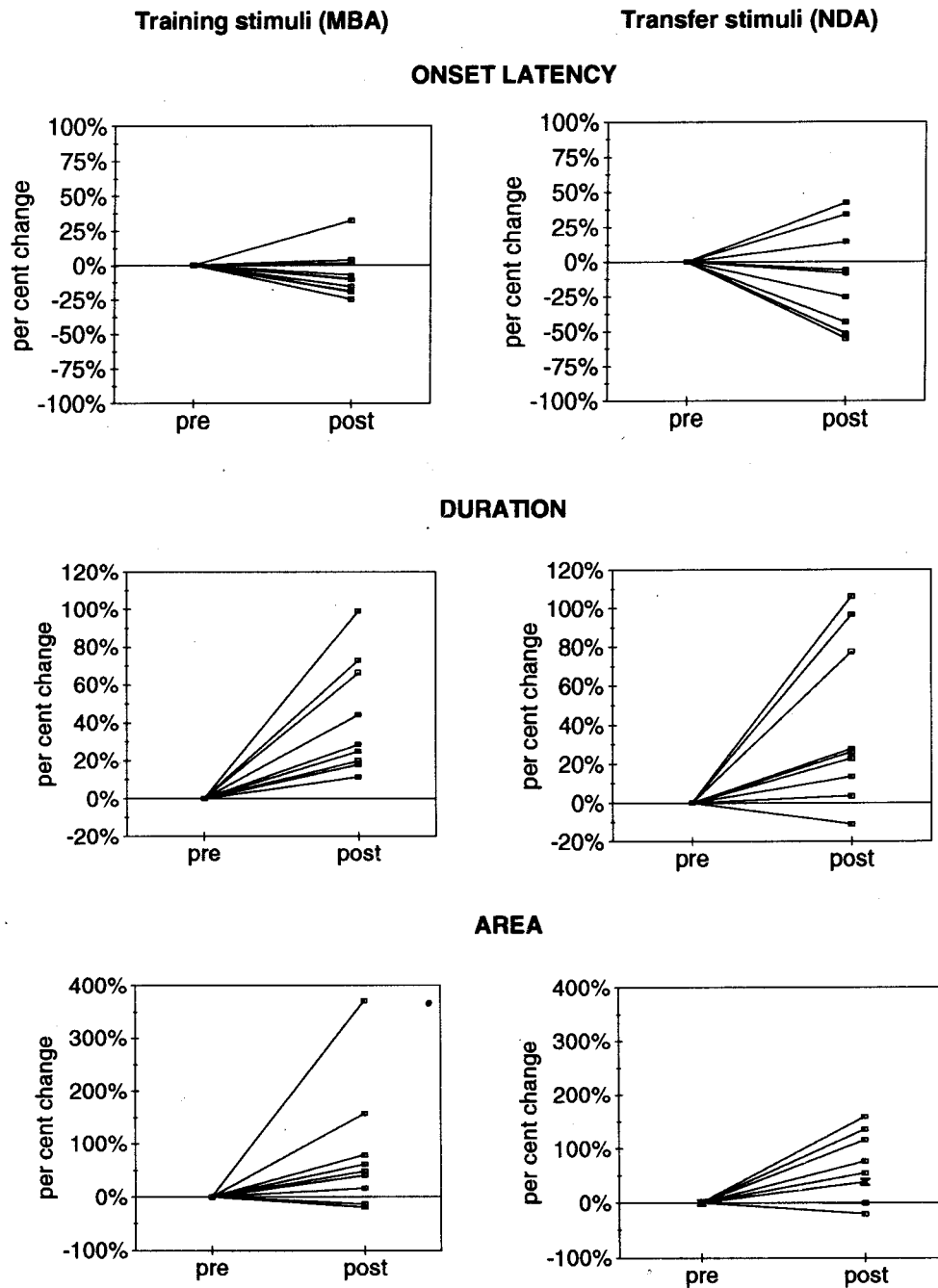


FIG. 7. Individual subject changes in MMN onset latency, duration, and area after training are shown for the experimental group. Posttraining changes are plotted relative to normalized pretraining values and show a significant increase in both duration and area as well as a significant decrease in onset latency. Note that overlapping individual scores are indicated by the X symbol.

may yield information pertaining to the site(s) of perceptual learning, and how learning manifests itself as behavior.

IV. CONCLUSIONS

Ultimately, the goal of successful (re)habilitative training is to have listeners transfer learned behaviors to new listening situations. The behavioral results of this study have confirmed that controlled laboratory-based training procedures can be effective in modifying listeners' preattentive phonetic perception in a short period of time and that the learned behavior transfers to new acoustic conditions. The

corresponding neurophysiologic findings are significant from a clinical perspective because they provide an objective approach for examining the neurophysiologic change that may occur following (re)habilitation efforts in populations such as hearing-impaired individuals and children with auditory-based learning problems, who may not be able to participate in behavioral testing. From a more theoretical standpoint, these findings establish an avenue for examining the human capacity for change, and for understanding neurophysiologic processes associated with foreign language training, musician ear training, and other forms of auditory learning.

ACKNOWLEDGMENTS

The authors thank Dawn Burton and Koch, Ph.D., Steve Zecker, Ph.D., Ann Bradlow, Ph.D., and Trent Nicol, the two JASA reviewers as well as the associate editor for their helpful comments. This work was supported by the National Institutes of Health R01 DC 01510.

- Ahissar, M., and Hochstein, S. (1993). "attentional control of early perceptual learning," *Proc. Natl. Acad. Sci. USA* **90**, 5718–5722.
- Alho, K., Woods, D. L., Algazi, A., and Näätänen, R. (1992). "Intermodal selective attention. II. Effects of attentional load on processing of auditory and visual stimuli in central space," *Electroencephalogr. Clin. Neurophysiol.* **85**, 356–368.
- Aslin, R. L., and Pisoni, D. B. (1980). "Some development processes in speech perception," in *Child Phonology: Perception and Production*, edited by G. Yeni-Komshian, J. F. Kavanagh, and C. A. Ferguson (Academic, New York), pp. 67–96.
- Bakin, J. S., and Weinberger, N. M. (1990). "Classical conditioning induces CS specific receptive field plasticity in the auditory cortex of the guinea pig," *Brain Res.* **536**, 271–286.
- Berger, T. W., Alger, B., and Thompson, R. F. (1976). "Neuronal substrate of classical conditioning in the hippocampus," *Science* **192**, 483–485.
- Blackburn, C. C., and Sachs, M. B. (1990). "The representations of the steady-state vowel sound /e/ in the discharge patterns of cat anteroventral cochlear nucleus neurons," *J. Neurophysiol.* **63**, 1191–1212.
- Buchwald, J. S., Halas, E. S., and Schramm, S. (1966). "Changes in cortical and subcortical unit activity during behavioral conditioning," *Physiol. Behav.* **1**, 11–22.
- Carney, A. E., Widin, G. P., and Viemeister, N. F. (1977). "Noncategorical perception of stop consonants differing in VOT," *J. Acoust. Soc. Am.* **62**, 961–970.
- Csèpe, V. (1995). "On the origin and development of the mismatch negativity," *Ear Hear.* **16**, 91–104.
- Csèpe, V., Karmos, G., and Molnar, M. (1987). "Evoked potential correlates of stimulus deviance during wakefulness and sleep in the cat-animal model of mismatched negativity," *Electroencephalogr. Clin. Neurophysiol.* **66**, 571–578.
- Delgutte, B., and Kiang, N. Y. S. (1984). "Speech coding in the auditory nerve. I. Vowel-like sounds," *J. Acoust. Soc. Am.* **75**, 866–878.
- Diamond, D. M., and Weinberger, N. M. (1984). "Physiologic plasticity of single neurons in auditory cortex of the cat during acquisition of the pupillary conditioned response. I: Secondary field (AII)," *Behav. Neurol.* **98**, 189–210.
- Diamond, D. M., and Weinberger, N. M. (1986). "Classical conditioning rapidly induced changes in frequency receptive fields of single neurons in secondary and ventral ectosylvian auditory cortical fields," *Brain Res.* **372**, 357–360.
- Diamond, N. M., and Weinberger, N. M. (1989). "Role of context in the expression of learning-induced plasticity of single neurons in auditory cortex," *Behav. Neurol.* **103**, 471–494.
- Disterhoft, J. F., and Stuart, D. K. (1976). "Trial sequence of changed unit activity in the auditory system of alert rat during conditioned response acquisition and extinction," *J. Neurophysiol.* **39**, 266–281.
- Disterhoft, J. F., and Stuart, D. K. (1977). "Differentiated short latency response increases after conditioning in inferior colliculus of alert rat," *Brain Res.* **130**, 315–334.
- Edeline, J. M., Neuenschwander-El Massioui, N., and Dutrieux, G. (1990). "Frequency-specific cellular changes in the auditory system during acquisition and reversal of discriminative conditioning," *Psychobiology* **18**, 382–393.
- Edeline, J. M., Pham, P., and Weinberger, N. M. (1993). "Rapid development of learning-induced receptive field plasticity in the auditory cortex," *Behav. Neurol.* **107**, 539–551.
- Eggmont, J. J. (1995). "Representation of a voice onset continuum in primary auditory cortex of the cat," *J. Acoust. Soc. Am.* **98**, 911–920.
- Giard, M., Perrin, F., Pernier, J., and Bouchet, P. (1990). "Brain generators implicated in the processing of auditory stimulus deviance: a topographic event-related potential study," *Psychophysiology* **27**, 627–640.
- Hari, R., Hamalainen, J., Ilmoniemi, R., Kaukoranta, E., Reinikainen, K., Salminen, J., Alho, K., Näätänen, R., and Sams, M. (1984). "Responses of the primary auditory cortex to pitch changes in a sequence of tone pips: neuromagnetic recordings in man," *Neurosci. Lett.* **50**, 127–132.
- Jamieson, D., and Morosan, D. (1986). "Training nonnative speech contrasts in adults: Acquisition of the English / /-θ/ contrast by francophones," *Percept. Psychophys.* **40**, 205–215.
- Jamieson, D., and Morosan, D. (1989). "Training new, nonnative speech contrasts in adults: A comparison of the prototype and perceptual fading techniques," *Can. J. Psychol.* **43**, 88–96.
- Jasper, H. H. (1958). "The ten-twenty system of the international federation," *Electroencephalogr. Clin. Neurophysiol.* **10**, 371–375.
- Javitt, D., Schroeder, C., Steinschneider, M., Arezzo, J., and Vaughan, Jr., H. (1992). "Demonstration of mismatch negativity in monkey," *Electroencephalogr. Clin. Neurophysiol.* **83**, 87–90.
- Kaas, J. H., Merzinech, M. M., and Killackey, H. P. (1983). "The reorganization of somatosensory cortex following peripheral nerve damage in adults and developing mammals," *Annu. Rev. Neurosci.* **6**, 325–356.
- Klatt, D. (1980). "Software for cascade/parallel formant synthesizer," *J. Acoust. Soc. Am.* **67**, 971–995.
- Kraus, N., and Disterhoft, J. F. (1982). "Response plasticity of single neurons in rabbit association cortex during tone-signalled learning," *Brain Res.* **246**, 205–215.
- Kraus, N., Micco, A., Koch, D., McGee, T., Carrell, T., Sharma, A., and Nicol, T. (1993). "Mismatch negativity in the neurophysiologic/behavioral evaluation of auditory processing deficits: A case study," *Ear Hear.* **14**, 223–234.
- Kraus, N., McGee, T., Littman, T., Nicol, T., and King, C. (1994a). "Encoding of acoustic change involves the non-primary auditory thalamus," *J. Neurophysiol.* **72**, 1270–1277.
- Kraus, N., McGee, T., Carrell, T., King, C., Littman, T., and Nicol, T. (1994b). "Discrimination of speech-like signals in auditory thalamus and cortex," *J. Acoust. Soc. Am.* **96**, 2758–2768.
- Kraus, N., McGee, T., Carrell, T., King, C., Tremblay, K., and Nicol, T. (1995a). "Central auditory system plasticity with speech discrimination training," *J. Cog. Neurosci.* **7**, 25–32.
- Kraus, N., McGee, T., Carrell, T., and Sharma, T. (1995b). "Neurophysiologic bases of speech discrimination," *Ear Hear.* **16**, 19–37.
- Kraus, N., McGee, T., Carrell, T., Zecker, S. G., Nicol, T., and Koch, D. (1996). "Neurophysiologic response reflect behavioral discrimination deficits in children with learning problems," *Science* **273**, 971–972.
- Lively, S., Logan, J., and Pisoni, D. (1993). "Training Japanese listeners to identify English /r/ and /l/: II. The role of phonetic environment and talker variability in learning new perceptual categories," *J. Acoust. Soc. Am.* **94**, 1242–1255.
- Lively, S., Pisoni, D., Yamada, R. A., Yoh'ichi, T., and Yamada, T. (1994). "Training Japanese listeners to identify English /r/ and /l/. III. Long-term retention of new phonetic categories," *J. Acoust. Soc. Am.* **96**, 2076–2087.
- Logan, J. S., Lively, S. E., and Pisoni, D. B. (1991). "Training Japanese listeners to identify English /r/ and /l/: A first report," *J. Acoust. Soc. Am.* **89**, 874–886.
- McClaskey, C., Pisoni, D., and Carrell, T. (1983). "Transfer of training of a new linguistic contrast in voicing," *Percept. Psychophys.* **34**, 323–330.
- McGee, T., Kraus, N., King, C., Nicol, T., and Carrell, T. (1996). "Acoustic elements of speech-like stimuli are reflected in surface recorded responses over the guinea pig temporal lobe," *J. Acoust. Soc. Am.* **99**, 3606–3614.
- McGee, T., Kraus, N., and Nicol, T. (1997a). "Is it really a mismatch negativity? an assessment of methods for determining response validity in individual subjects," *Electroenceph. Clin. Neurophysiol.* **104**, 359–368.
- McGee, T., Nicol, T., and Kraus, N. (1997b). "Improving the MMN signal-to-noise ratio." XV International Evoked Response Audiometry Study Group (IERASG) Abstract, p. 49.
- Merzinech, M. M., and Jenkins, W. M. (1983). "Dynamic maintenance and alterability of cortical maps in adults; some implications," in *Hearing—Physiological Bases and Psychophysics*, edited by R. Klinke and R. Hartmann (Oxford U.P., New York), pp. 162–167.
- Merzinech, M. M., Kaas, J. H., Wall, J. T., Sur, M., and Nelson, R. J. (1983). "Progression of change following median nerve section in the cortical representation of the hand in areas 3b and 1 in adult owl and squirrel monkey," *Neuroscience (NY)* **10**, 639–665.
- Merzinech, M. M., Nelson, R. J., Stryker, M. P., Cynader, M. S., Schoppmann, A., and Zook, J. M. (1984). "Somatosensory cortical map changes following digit amputation in adult monkeys," *J. Comp. Neurol.* **224**, 591–605.
- Micco, A., Kraus, N., Koch, D., McGee, T., Carrell, T., Sharma, A., Nicol, T., and Wiet, R. (1995). "Speech-evoked cognitive potentials in cochlear implant recipients," *Am. J. Otolaryngol.* **16**, 514–520.

- Näätänen, R., and Picton, T. (1987). "The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure," *Psychophysiology* **24**, 375–425.
- Näätänen, R., Gaillard, A., and Mantysalo, S. (1978). "Early selective attention effect on evoked potential reinterpreted," *Acta Psychol.* **42**, 313–329.
- Näätänen, R., Gaillard, A., and Mantysalo, S. (1980). "Brain potential correlated in voluntary and involuntary attention," *Prog. Brain Res.* **54**, 343–348.
- Näätänen, R., Schroger, E., Karakas, S., Tervaniemi, M., and Paavilainen, P. (1993). "Development of neural representations for complex sound patterns in the human brain," *NeuroReport* **4**, 503–506.
- Olds, J., Disterhoft, J. T., Segal, M., Kornblith, C. L., and Hirsh, R. (1972). "Learning centers of rat brain mapped by measuring the latencies of conditioned unit responses," *J. Neurophysiol.* **35**, 202–219.
- Paavilainen, P., Alho, K., Reinikainen, K., Sams, M., and Näätänen, R. (1972). "Right hemisphere dominance of different mismatch negativities," *Electroencephalogr. Clin. Neurophysiol.* **78**, 466–479.
- Pisoni, D. B., Aslin, R. N., Perey, A. J., and Hennessy, B. L. (1982). "Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants," *J. Exp. Psychol.* **8**, 297–314.
- Recanzone, G., Schreiner, C., and Merzenich, M. (1993). "Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys," *J. Neurosci.* **13**, 87–104.
- Ryugo, D. K., and Weinberger, N. M. (1978). "Differential plasticity of morphologically distinct neuron populations in the medial geniculate body of the cat during classical conditioning," *Behav. Biol.* **22**, 275–301.
- Sams, M., and Näätänen, R. (1991). "Neuromagnetic responses of the human auditory cortex to short frequency glides," *Neurosci. Lett.* **121**, 43–46.
- Sams, M., Paavilainen, P., Alho, K., and Näätänen, R. (1985). "Auditory frequency discrimination and event-related potentials," *Electroencephalogr. Clin. Neurophysiol.* **62**, 437–448.
- Scherg, M., and Picton, T. (1990). "Brain electric source analysis of mismatch negativity," in *Psychophysiological Brain Research*, edited by C. Brunia, A. Gaillard, and A. Kok (Tilberg U. P., Tilberg), pp. 94–98.
- Sharma, A., and Kraus, N. (1995). "Effect of contextual variations in pitch and phonetic processing: Neurophysiologic correlates." *Assoc. for Research in Otolaryngology. Eighteenth Midwinter Research Meeting, Abstract #729.*
- Sinex, D. G., and McDonald, L. P. (1988). "Average discharge rate representation of voice onset time in the chinchilla auditory nerve," *J. Acoust. Soc. Am.* **83**, 1817–1827.
- Steinschneider, M., Arezzo, J., and Vaughan, Jr., H. (1982). "Speech-evoked activity in the auditory radiations and cortex of the awake monkey," *Brain Res.* **252**, 353–365.
- Steinschneider, M., Arezzo, J., and Vaughan, Jr., H. (1990). "Tonotopic features of speech-evoked activity in primate auditory cortex," *Brain Res.* **519**, 158–168.
- Steinschneider, M., Schroeder, C., Arezzo, J., and Vaughan, Jr., H. (1995). "Physiological correlates of the voice onset time boundary in primary auditory cortex (AI) of the awake monkey: temporal response patterns," *Brain Lang.* **48**, 326–340.
- Strange, W., and Dittmann, S. (1984). "Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English," *Percept. Psychophys.* **36**, 131–145.
- Weinberger, N. M. (1993). "Learning-induced changes of auditory receptive fields [Review]," *Curr. Opin. Neurobiol.* **3**(4), 570–577.
- Weinberger, N. M., Hopkins, W., and Diamond, D. M. (1984). "Physiologic plasticity of single neurons in auditory cortex of the cat during acquisition of the pupillary conditioned response. I: Primary field (AI)," *Behav. Neurosci.* **98**, 171–188.
- Werker, J., and Tees, R. (1984). "Phonemic and phonetic factors in adult cross-language speech perception," *J. Acoust. Soc. Am.* **75**, 1866–1878.